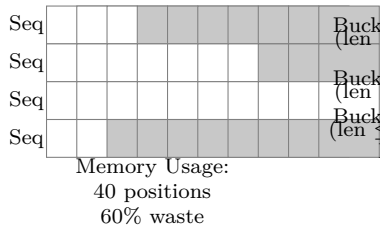
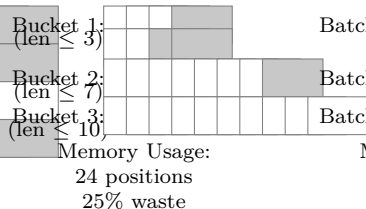


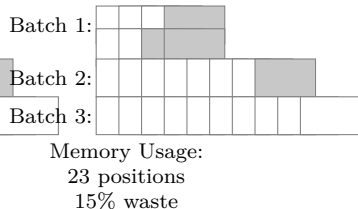
1. Naive Padding



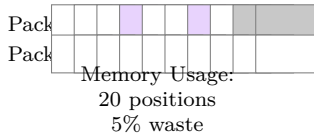
2. Length Bucketing



3. Dynamic Batching



4. Packed Sequences



5. Variable Length (Future)



Performance Comparison

Efficiency

Naive

Bucketing

Dynamic

Packed

Variable

Real Token
Padding
Separator

Strategy